1. Project Name

   Prediction of the Helpfulness of a Review among Amazon Product Reviews

2. Description of Problem

   As an avid user of Amazon purchasing almost everything from pantry to video, I am always in need of helpful reviews of a product while comparing multiple at a time. There are many studies on the web which analyze whether a review on a product is positive or not, however, there are few studies on which types of reviews are found helpful to the prospective buyers on Amazon. If what affects the helpfulness of a review is analyzed, it will be possible for Amazon to give suggestions for its users when they are reviewing a product, and this will also help Amazon to easily categorize and remove not helpful or spam reviews.

3. Project Clients

   This project's audience is considered to be Amazon, or any Amazon user, who would like to determine the useful reviews and only skim them to reach a conclusion on a product he or she would like to purchase.

4. Data to be used

   The dataset is provided on different product categories by Julian McAuley, through its website http://jmcauley.ucsd.edu/data/amazon/links.html.

5. Outline of the Approach

   After the data are acquired, it is aimed to clean and explore the data. Since the data is from different categories ranging from books to instant video, a small subset of one category will be considered for analysis and model building first. Later, the analysis will be applied to all categories which contain at least one helpful review as it is given in the dataset.

6. Deliverables

   A short report summarizing the description of the project and dataset, along with important findings and results will be delivered.